A Linear Scaling Subspace Iteration Algorithm with Optimally Localized Non-Orthogonal Wave Functions for Kohn-Sham Density Functional Theory

C.J. García-Cervera
Mathematics Department, University of California, Santa Barbara, CA 93106

J. Lu and Y. Xuan
PACM, Princeton University, Princeton, NJ 08540

W. E
Department of Mathematics and PACM, Princeton University, Princeton, NJ 08540

(Dated: October 3, 2008)

We present a new linear scaling method for electronic structure computations in the context of Kohn-Sham density functional theory (DFT). The method is based on a subspace iteration, and takes advantage of the non-orthogonal formulation of the Kohn-Sham functional, and the improved localization properties of non-orthogonal wave functions. A one-dimensional linear problem is introduced as a benchmark for the analysis of linear scaling algorithms for Kohn-Sham DFT. Using this one-dimensional model, we study the convergence properties of the localized subspace iteration algorithm presented. We demonstrate the efficiency of the algorithm for practical applications by performing fully three-dimensional computations of the electronic density of alkane chains.

PACS numbers:

I. INTRODUCTION

Kohn-Sham density functional theory (DFT) is a very popular tool for electronic structure analysis, with applications in materials science, chemistry, and other areas. Compared with the quantum many-body problem, which describes a system with \( N \) electrons by a \( 3N \)-dimensional antisymmetric wave function (ignoring spin degeneracy), the Kohn-Sham density functional theory is an approximate model which describes such a system by \( N \) one-particle wave functions. By suitably choosing the functionals, Kohn-Sham DFT can become a very effective alternative to the quantum many-body problem, with satisfactory accuracy and much improved efficiency [1].

The Kohn-Sham equations are a system of nonlinear eigenvalue problems. The traditional self-consistent approach [2] for the solution of this eigenvalue problem consists of two nested iterations: in the inner iteration, the orbitals \( \{ \psi_j \}^N_{j=1} \) are obtained by a process of diagonalization and orthogonalization; in the outer iteration, the electron density is updated until self-consistency is reached. The diagonalization and/or orthogonalization procedure scales typically as \( O(N^3) \), which is prohibitively expensive for relatively small problems.

In the past twenty years, a number of new methodologies that exploit the locality of the problem in order to reduce the computational complexity, have appeared in the literature [10, 20]. Locality, in quantum mechanics, refers to the property that a small disturbance in a molecule only has a local effect in the electron density, a phenomenon coined by W. Kohn as nearsightedness [3]. Examples of these attempts are the Divide and Conquer methods [4, 5], Orbital Minimization [6], Density Matrix Minimization [7, 8], and the Fermi Operator Expansion [9]. For a review of these ideas and other methodologies, see [10].

In this article, we describe in detail the linear scaling algorithm introduced in [11]. The algorithm is based on a subspace iteration procedure with localized non-orthogonal wave functions. The general ideas of such a localized subspace iteration (LSI) were first introduced in [12]. In the context of electronic structure analysis, the subspace iteration approach has been used in [13]. Our algorithm replaces the diagonalization and orthogonalization step in the subspace iteration algorithm introduced in [13] by a localization procedure (described in detail in section II B). Linear scaling is achieved in our algorithm by exploiting the localization properties of non-orthogonal wave functions.

Kohn-Sham DFT and the Kohn-Sham equations are fairly complicated problems. As a variational problem, the Kohn-Sham functional is nonlinear, nonlocal and non-convex. The Kohn-Sham equations are a set of nonlinear and nonlocal eigenvalue problems. For practical applications, there is the added complexity coming from the exchange and correlation functional and the pseudopotential. These factors together make Kohn-Sham DFT a rather formidable problem from a mathematical and numerical viewpoint. However, some of these complexities are inessential for the purpose of developing and understanding numerical algorithms. For these reasons, we introduce here a one-dimensional linear model and use it as a benchmark for clarifying the essential fea-
tures of the localized subspace iteration algorithm, its performance and its limitations. This one-dimensional problem, as well as its nonlinear version, which will be discussed in subsequent publications, is very simple, and it has many of the essential features of the Kohn-Sham DFT. In subsequent papers, we will present, using this model, a systematic analysis of other linear scaling algorithms and self-consistent iterations for Kohn-Sham DFT.

One interesting feature of the LSI is that it does not converge to a fixed point: Under favorable conditions (reasonable gap size, large enough localization region, etc), the error will decrease to a rather small value and then start to fluctuate. This is due to the conflicting effect of the filtering and truncation steps in the LSI, and should be a general feature of linear scaling algorithms involving truncation. Using the one-dimensional model, we are able to understand the convergence properties of the LSI in considerable detail.

This paper is organized as follows: In the remainder of this introduction, we will review Kohn-Sham DFT and some general numerical issues. The linear scaling algorithm and its main components are described in considerable detail in section [11]. We introduce the aforementioned one-dimensional model problem, and analyze how the different components of the algorithm impact its performance and accuracy, in section [III]. The fully three-dimensional implementation of the algorithm for Kohn-Sham DFT is described in section [IV] where the results are validated, and linear scaling is demonstrated for practical examples.

A. Kohn-Sham density functional theory

Consider a system consisting of \(N_a\) atoms and \(2N\) electrons. In Kohn-Sham DFT, the state of the system is described by a set of \(N\) wave functions \(\{\psi_j\}_{j=1}^N\), representing the interacting electrons. In the Born-Oppenheimer approximation, the spinless Kohn-Sham energy functional can be written as [2, 14]

\[
E_{\text{KS}}[\{\psi_j\}] = 2 \sum_{j=1}^N \int_{\mathbb{R}^3} \psi_j \left( -\frac{1}{2} \Delta \psi_j \right) \, dx + E_{\text{XC}}[\rho] \\
+ \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{(\rho - m)(x)(\rho - m)(y)}{|x - y|} \, dx \, dy + E_{\text{PS}}[\{\psi_i\}],
\]

(1)

where the electron density is defined as

\[
\rho(x) = 2 \sum_{j=1}^N |\psi_j(x)|^2,
\]

and the wave functions \(\{\psi_i\}_{i=1}^N\) are required to be orthonormal:

\[
\int_{\mathbb{R}^3} \psi_i(x) \psi_j(x) \, dx = \delta_{ij}, \quad i, j = 1, \ldots, N.
\]

(2)

In [1], the exact form of the exchange and correlation functional \(E_{\text{XC}}[\rho]\) is unknown, and needs to be approximated. We adopt here the local density approximation (LDA), in which this term is written as

\[
E_{\text{XC}}[\rho] = \int_{\mathbb{R}^3} \rho \varepsilon(\rho),
\]

(3)

where \(\varepsilon\) is a function of the density alone. The generalized gradient approximation (GGA) [15], in which \(\varepsilon\) depends on the gradient of \(\rho\) as well, can be used instead, and this will not change the discussions below in any significant way.

We consider \(\rho\) to be the density of valence electrons, and their electrostatic interactions with the nuclei and the core electrons are described by a so called pseudopotential. The pseudopotential is completely defined by an ionic function, \(m\), and an ionic pseudopotential operator, \(\tilde{V}_{PS}\). The ionic function is

\[
m(x) = \sum_{j=1}^{N_a} m^a(x - R_j),
\]

where \(m^a\) is a localized function (usually with exponential decay), and \(\{R_j\}_{j=1}^{N_a}\) are the locations of the atoms. In this article, we use the norm conserving Troullier-Martins pseudopotential [16], in the Kleinman-Bylander form [17]. We therefore define the ionic pseudopotential operator as

\[
\tilde{V}_{PS}\psi(x) = \sum_{j=1}^{N_a} \left( \tilde{V}_{local}^j (x - R_j) \psi(x) \\
+ \sum_{l=0}^{l_{\text{max}}} \sum_{m=-l}^{l} \int_{\mathbb{R}^3} \beta^l_{lm}(y - R_j) \psi(y) \, dy \, \beta^l_{lm}(x - R_j) \right),
\]

(4)

where \(\tilde{V}_{local}^j\) is the local potential for the \(j\)-th atom (by which we mean that it acts by pointwise multiplication), and \(\{\beta^l_{lm}\}\) projects \(\psi\) onto the subspace generated by the corresponding pseudo-wave function of the \(j\)-th atom. The indices \(\{l, m\}\) indicate the angular momentum components of the pseudo-wave function. By introducing the ionic function \(m\), we can make sure that \(\tilde{V}_{local}^j\) and \(\{\beta^l_{lm}\}\) are all compactly supported. This is important in order to achieve linear scaling.

In [1], the pseudopotential energy \(E_{PS}\) is defined as

\[
E_{PS}[\{\psi_i\}] = 2 \sum_{i=1}^N \int_{\mathbb{R}^3} \psi_i(x) \tilde{V}_{PS} \psi_i(x) \, dx.
\]

(5)

When the Kohn-Sham functional [1] is minimized under the orthogonality constraint [2], the Euler-Lagrange equations lead to the following nonlinear eigenvalue problem:

\[
H \psi_i = \varepsilon_i \psi_i; \quad i = 1, \ldots, N.
\]

(6)
The Hamiltonian in \([6]\) is defined as

\[
H = -\frac{1}{2} \Delta + V_{\text{eff}}[\rho], \tag{7}
\]

where

\[
V_{\text{eff}}[\rho](x) = \int_{\mathbb{R}^3} \frac{(\rho - m)(y)}{|x - y|} dy + V_{\text{PS}}(x) + \frac{\delta E_{\text{XC}}[\rho]}{\delta \rho}.
\]

The starting point of our linear scaling approach is the reformulation of the Kohn-Sham problem in terms of non-orthogonal wave functions, as introduced in \([6]\). Given \(N\) linearly independent wave functions, \(\{\psi_j\}_{j=1}^N\), define the overlap or Gram matrix as

\[
S_{jk} = \int_{\mathbb{R}^3} \psi_j(x)\psi_k(x), \quad j, k = 1, \ldots, N. \tag{8}
\]

Then,

\[
E_{\text{KS}}[\{\psi_j\}] = 2 \sum_{j,k} (S^{-1})_{jk} \int_{\mathbb{R}^3} \psi_j \left( -\frac{1}{2} \Delta \psi_k \right) dx \\
+ E_{\text{XC}}[\rho] + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{(\rho - m)(x)(\rho - m)(y)}{|x - y|} dxdy \\
+ E_{\text{PS}}[\{\psi_i\}], \tag{9}
\]

where the electron density is now defined by

\[
\rho(x) = 2 \sum_{j,k} \psi_j(x)(S^{-1})_{jk} \psi_k(x), \tag{10}
\]

and the pseudopotential energy is now

\[
E_{\text{PS}}[\{\psi_i\}] = 2 \sum_{j,k} (S^{-1})_{jk} \int_{\mathbb{R}^3} \psi_j(x)V_{\text{PS}} \psi_k(x) dx. \tag{11}
\]

It is easy to show that the Kohn-Sham functional \([6]\) is invariant under nonsingular linear transformations of the \(N\)-dimensional subspace spanned by the wave functions:

\[
\text{span}\{\psi_i\} = \text{span}\{\phi_i\} \implies E_{\text{KS}}[\{\psi_i\}] = E_{\text{KS}}[\{\phi_i\}]. \tag{12}
\]

As a consequence, the Kohn-Sham functional can be thought of as a functional acting on subspaces, and the goal is to find the \textit{occupied subspace}. This is the subspace generated by the eigenfunctions corresponding to the smallest eigenvalues of the self-consistent Hamiltonian. The advantage of this viewpoint is that the specific representation of the subspace is not relevant, and therefore we can choose the representation that is most convenient for our purposes. Linear scaling can be achieved by choosing a representation in terms of optimally localized non-orthogonal wave functions, as described in \([12]\). Non-orthogonal wave functions have better localization properties than orthogonal orbitals \([18, 19]\), which constitutes an additional advantage of this approach.

\[\text{B. Numerical algorithms}\]

A numerical algorithm for the Kohn-Sham problem should have the following basic components:

1. A discretization procedure that translates the continuous problem into a discrete (finite dimensional) problem. Typical discretization procedures include the plane wave methods, methods based on a linear combination of atomic orbitals, finite difference and finite element methods \([20]\).

2. An algorithm for handling the linearized finite dimensional problem. In particular, assuming that the electron density is given, we need to find the optimal eigen-subspace for the Kohn-Sham operator with a known potential. This can be done by diagonalization, or some other procedure \([11, 20]\).

3. A nonlinear iteration procedure to achieve self-consistency. One can use either a mixing procedure or some self-consistent iteration \([20]\).

In the present paper, we will focus on the second component, i.e., finding the minimizing subspace of the Kohn-Sham operator. For the numerical results presented below, we use a simple second order finite difference discretization and simple linear mixing:

\[
\rho^{\text{new}} = \alpha \rho^{\text{old}} + (1 - \alpha) \rho^*, \tag{13}
\]

where \(\rho^*\) is the outcome of component 2 above, and \(\alpha \in [0,1]\). These are chosen for their simplicity. There is no difficulty in using higher order finite difference or finite element discretizations or other mixing schemes. However, since we will focus on localized wave functions, our work does not apply to the case when plane waves are used.

\[\text{II. LOCALIZED SUBSPACE ITERATION FOR KNOHN-SHAM DENSITY FUNCTIONAL THEORY}\]

\[\text{A. Subspace iteration}\]

Consider the problem of computing the largest eigenvalue (in magnitude) of a symmetric matrix \(H\). For simplicity of presentation, let us assume that the eigenvalues of the matrix \(H\) are \(0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N < \infty\). An effective approach is the power method \([21]\):

1. Given an initial vector \(v^0\).

2. For \(k \geq 0\), define

\[
(a) \quad v^{k+1} = \frac{Hv^k}{\|Hv^k\|}, \tag{14}
\]

\[
(b) \quad \mu^{k+1} = (v^{k+1})^T \cdot Hv^{k+1}.
\]
3. Repeat until $|\mu^{k+1} - \mu^k| \leq \text{Tolerance}$.

One can show that the power method converges, i.e., $\mu^k \to \lambda_N$, with convergence rate

$$\tau = \frac{\lambda_{N-1}}{\lambda_N} < 1.$$ (15)

If other eigenvalues are required, one can use the shifted power method, or the inverse shifted power method, which are essentially the power method applied to $H - \lambda I$ and $(H - \lambda I)^{-1}$, respectively.

The subspace iteration is a straightforward generalization of the power method, applied to a subspace. A typical procedure is the following:

1. Given an initial space $V_0$ of dimension $M < N$, for each $k \geq 1$:
   (a) Calculate $W_k = HV_k$.
   (b) Orthogonalize the basis (QR decomposition, for example): $W_k = Q_k R_k$.
   (c) Let $V_k = Q_k$.

2. Repeat until convergence.

The orthogonalization step is necessary in order to ensure the linear independence of the vectors in the new space. In the context of electronic structure analysis, subspace iteration has been used by Zhou, Saad, Tiago, and Chelikowsky [13].

It can be proved that the previous algorithm converges, and that the rate of convergence is [21]

$$\tau = \frac{\lambda_M}{\lambda_{M+1}} < 1.$$ (16)

The cost of the orthogonalization step scales like $O(N^3)$. A key idea in the new algorithm that we present here is to replace this step by a localization procedure: Given a candidate subspace, we find an optimally localized basis for this subspace. The details of this procedure are explained below.

B. Localization

Localization has been recognized by several authors as one of the key ideas in achieving linear scaling [3, 8, 22]. One of the first implementations of locality in Kohn-Sham DFT consisted on computing the so-called Maximally Localized Wannier Functions [22] in periodic systems, and for the orthogonal formulation of Kohn-Sham DFT [1].

This concept was generalized to the non-orthogonal case in [12]. Given a linear space $V = \text{span} \{\psi_j\}_{j=1}^N$ of dimension $N$, and a given smooth weight function $w \geq 0$, the maximally localized non-orthogonal wave function $\tilde{\psi}$ is defined as

$$\tilde{\psi} = \arg \min_{\phi \in V, \| \phi \|=1} \int_{\mathbb{R}^3} w(x)|\phi(x)|^2 dx.$$ (17)

Non-orthogonal wave functions have better localization properties than orthogonal Wannier functions. The improved decay properties of these wave functions can be found in [19], and are summarized in the following theorem, borrowed from the same source:

**Theorem II.1** Suppose $p$ is the index of the first nonzero term in $\{w^{(2p)}(0)\}$, $p \geq 1$, then the generalized non-orthogonal Wannier function for an insulator decays as $|x - x_c|^{-\left(2p+1/2\right)} \exp(-\hbar |x - x_c|)$.

Given a set of wave functions, $\{\psi_j\}_{j=1}^N$, centered at the locations $\{b_j\}_{j=1}^N$, respectively, we obtain a localized representation of $V = \text{span} \{\psi_j\}_{j=1}^N$ by minimizing, for each $j \in \{1, \ldots, N\}$,

$$F[\phi] = \int_{\mathbb{R}^3} |x - b_j|^{2p} |\phi(x)|^2 dx,$$ (18)

among functions $\phi$ of the form

$$\phi(x) = \sum_{k=1}^r \alpha_k \psi_k(x) \in V.$$ (19)

The minimization (18) leads to $W a = \lambda S a$, where

$$W_{kl} = \int_{\mathbb{R}^3} |x - b_j|^{2p} \psi_k(x) \psi_l(x) dx, \quad k, l = 1, \ldots, r,$$

and $\lambda$ is the smallest generalized eigenvalue.

Given that the wave functions $\{\psi_j\}_{j=1}^N$ are compactly supported, only a fixed number of wave functions appear in (19), i.e., $r$ is bounded independently of $N$. Therefore the localized basis can be obtained with $O(N)$ scaling. It is important to notice that the new, localized wave functions, span the same space as the one we started with. Since the new localized wave functions span the same subspace as the original ones, their Kohn-Sham energy is the same.

There are several additional important technical components of our localized subspace iteration algorithm that need to be explained:

1. The filtering procedure, which is an improvement over the power method.
2. The computation of the Fermi energy, necessary to define an effective filter, and
3. The computation of $S^{-1}$.

We now explain these components in detail.

C. Filtering

The power method, and its variants, can be understood as a simple filtering strategy, by which a certain portion of the spectrum of $H$ is amplified more than others. In view of its convergence rate, [16], it is clear that one can
construct a more efficient filter by applying a subspace iteration to \( P(H) \), where \( P \) is a polynomial of degree \( n \). Assuming that the polynomial \( P \) splits the spectrum of \( H \), in the sense that
\[
P(\lambda_i) \leq P(\lambda_M), \quad i = 1, \ldots, M, \tag{20}
\]
\[
P(\lambda_j) \geq P(\lambda_{M+1}), \quad j = M + 2, \ldots, M, \tag{21}
\]
the rate of convergence of the polynomial filtered subspace iteration is \( \kappa = \left| \frac{P(\lambda_M)}{P(\lambda_{M+1})} \right| \). \( \tag{22} \)

By a judicious choice of \( P \), we can get \( \kappa \ll \tau \) (see \( \PageIndex{16} \)), thereby accelerating the convergence of the subspace iteration considerably. It can be shown that the optimal choice for \( P \) is the Chebyshev polynomial of degree \( n \), \( \tau_n \).

The Chebyshev polynomial of degree \( n \) is defined as
\[
T_n(x) = \begin{cases} 
\cos(n \cos^{-1} x) & \text{if } |x| \leq 1, \\
(-1)^n \cos(n \cos^{-1} |x|) & \text{if } |x| \geq 1.
\end{cases} \tag{23}
\]

The Chebyshev polynomials have the property that \( |T_n(x)| \leq 1 \) for \( x \in [-1,1] \) and \( |T_n(x)| \gg 1 \) for \( |x| > 1 \). In addition, they satisfy the three-term recursion
\[
T_0(x) = 1; \quad T_1(x) = x; \quad T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x), \quad k \geq 1, \tag{24}
\]
which can be used to evaluate \( T_n(H) \) efficiently without explicitly computing the operator. The Chebyshev polynomial \( T_{10} \) is plotted in figure 1 to illustrate the fast growth outside \([-1,1]\).

To use the Chebyshev polynomial as an effective filter, the spectrum of the Hamiltonian is divided into \( \varepsilon_F < \varepsilon < \varepsilon_F \). The map \( x \rightarrow 2 \frac{x - \varepsilon_F}{\varepsilon_{\text{max}} - \varepsilon_F} - 1 \) transforms the interval \([\varepsilon_F, \varepsilon_{\text{max}}]\) into \([-1,1]\), and the interval \([\varepsilon, \varepsilon_F]\) is mapped into a subinterval of \((-\infty, -1]\). Therefore the rescaled Chebyshev polynomial
\[
T_n \left( 2 \frac{x - \varepsilon_F}{\varepsilon_{\text{max}} - \varepsilon_F} - 1 \right)
\]
amplifies the lower end of the spectrum, while leaving the high end unchanged in magnitude, effectively filtering out the high end of the spectrum.

\section{Computation of the Fermi Energy}

The Fermi energy \( \varepsilon_F \) is necessary in our algorithm in order to define an effective Chebyshev filter, as described in section \( \PageIndex{14} \). Given that we have no knowledge of the spectrum, this is not as straightforward as in codes that involve diagonalization \( \PageIndex{14} \). We estimate the Fermi energy by the maximum eigenvalue of the Ritz matrix. This eigenvalue can be approximated with linear complexity in the following way: Given a basis \( \Psi \) of the subspace \( V \), we know that \( \Psi S^{-1/2} \) is an orthonormal basis, where the Gram matrix \( S \) is defined in \( \PageIndex{8} \). The Ritz matrix \( R \in \mathbb{R}^{N \times N} \) is defined as
\[
R = (\Psi S^{-1/2})^T H \Psi S^{-1/2} = S^{-1/2} \Psi^T H \Psi S^{-1/2}. \tag{25}
\]
We are interested in the eigenvalues of \( R \), which are the same as the eigenvalues of \( S^{1/2} R S^{-1/2} = \Psi^T H \Psi S^{-1} \). The Fermi energy is then approximated as the maximum eigenvalue of \( R \), which can be approximated with a few steps of the power method. Note that the power method will converge to the maximum eigenvalue in magnitude. Therefore, to approximate the Fermi energy, we compute both the maximum and minimum eigenvalues in magnitude of the Ritz matrix, respectively \( \lambda_M \) and \( \lambda_m \). Once the maximum eigenvalue is approximated, the shifted power method can be used to obtain the minimum eigenvalue (in magnitude). Note that in the application of the power method, we do not need \( S^{-1} \) explicitly, but only \( S^{-1} v \) for some vector \( v \). This can be obtained by solving the system of equations \( S y = v \). Assuming that the condition number of \( S \) is bounded independently of \( N \), and given that \( S \) is sparse, the system of equations can be solved with the preconditioned conjugate gradient, using the main tridiagonal part of \( S \) as a preconditioner \( \PageIndex{24} \). Since the basis \( \Psi \) is localized, the Ritz matrix is sparse, and the required matrix-vector multiplications can be performed with linear scaling.

\section{Computation of \( S^{-1} \)}

The inverse of the Gram matrix \( S \), defined in \( \PageIndex{8} \), plays a fundamental role in the non-orthogonal formulation of Kohn-Sham DFT, as it appears in the energy functional itself \( \PageIndex{9} \), the definition of the electron density \( \PageIndex{10} \), and the pseudopotential \( \PageIndex{11} \). A direct computation of \( S^{-1} \) scales as \( O(N^3) \). However, under the assumption that both \( S \) and \( S^{-1} \) decay exponentially fast away from the diagonal, \( S^{-1} \) can be computed with \( O(N) \) complexity.
We have implemented the scaled third-order Newton-Schulz iteration described in [23] for the inverse matrix square root. Given the Gram matrix $S$, the iteration solves equation

$$SZS = I,$$

(26)

to compute $Z = S^{-1/2}$ in the following way:

1. Let $\lambda = \frac{1}{\|S\|_1}$, $Z_0 = I$, $Y_0 = S$ and $k = 0$;
2. Set $X_k = \lambda Y_k Z_k$, $T_k = \frac{1}{2}(15I - 10X_k + 3X_k^2)$, $Z_{k+1} = Z_k T_k$ and $Y_{k+1} = T_k Y_k$.
3. Repeat the previous step until $\|Z_{k+1} - Z_k\|_1 \leq$ Tolerance.
4. $S^{-1/2} = \lambda^{1/2} Z_k$ and $S^{-1} = \lambda Z_k^2$.

All operations are carried out using sparse matrix computations. The truncation of the successive matrices is done using the algorithm described in [25]. The main idea in this approach is to control the $l_1$-norm error at each step by discarding the blocks whose contribution to the next iterate are below a given tolerance.

F. LSI Algorithm for Kohn-Sham DFT

The algorithm for Kohn-Sham DFT, based on a Chebyshev-filtered subspace iteration with optimally localized non-orthogonal wave functions is [11]:

\begin{algorithm}
\textbf{Algorithm 1} : Localized Subspace Iteration for Kohn-Sham DFT.
1: Given (localized) wave functions $\Psi_0$.
2: repeat \{Self-Consistency Loop (SCF)\}
3: Compute the electron density, $\rho$.
4: Compute the effective potential, $V_{\text{eff}}[\rho]$.
5: repeat \{Localized Subspace Iteration\}
6: Filtering Step: $\Psi = T_n(H)\Psi$.
7: Localization Step: Localize $\psi_r$, for $r = 1, \cdots, N$.
8: Truncate beyond the cut-off radius.
9: until Convergence of the linear iteration.
10: Update the electron density (mixing).
11: until Convergence of the self-consistent iteration.
\end{algorithm}

III. PERFORMANCE EVALUATION USING A ONE-DIMENSIONAL MODEL

The original Kohn-Sham problem is a rather complicated problem in three dimensions. In order to better understand the essential features of the proposed algorithm, we will first study its behavior on a much simplified one-dimensional model. This one-dimensional model shares the most important features of a linearized Kohn-Sham model. It can be used as a testbed for systematically analyzing the accuracy and convergence of DFT algorithms.

A. The one-dimensional model

We consider an infinite array of atoms on a line with unit spacing: $X_i = i$, for $i \in \mathbb{Z}$. Each atom has one valence electron and we ignore spin degeneracy. The electrons are non-interacting, so that the electronic structure of the system is determined by solving linear eigenvalue problems (instead of nonlinear eigenvalue problems as in the full Kohn-Sham case):

$$H\psi_i = \epsilon_i \psi_i,$$

(27)

where the Hamiltonian is given by

$$H = -\frac{1}{2} \frac{\partial^2}{\partial x^2} + V(x).$$

(28)

The effective potential $V$ is a sum of Gaussian wells located at the atom sites,

$$V(x) = -\sum_{i \in \mathbb{Z}} \frac{a}{\sqrt{2\pi}\sigma^2} \exp(-(x - X_i)^2/2\sigma^2).$$

(29)

This model has two parameters: $a$, which characterizes the depth of the wells, and $\sigma$, which characterizes its width.

The band structure of this model is shown in figure 3 for several sets of parameter values. The first band is fully occupied and the second band is empty. The band gap $E_{\text{gap}}$ is the difference between the highest eigenvalue in the first band and the lowest eigenvalue in the second band (both obtained at the edge of the first Brillouin zone $k = \pi$ [25]). We are also interested in the ratio between the band gap and the width of the first band $E_{\text{width}}$. The relation between the band gap and the parameters $a$ and $\sigma$ of the potential is shown in figure 3. It is observed that the gap is proportional to $\sqrt{a}/\sigma$. By changing parameters, we may change the model from a well gapped insulator to a metal-like system. We are going to test the performance of LSI algorithms in different situations.

B. Convergence issue of the LSI and related algorithms

Recall that each LSI iteration contains three steps: Filtering, localization and truncation. Starting from a given subspace, the filtering step makes the subspace closer to the occupied subspace by filtering out the higher spectrum. The localization step finds a better representation while keeping the subspace unchanged. It is clear that without truncation, the LSI iteration will converge and will give the correct occupied subspace of the Hamiltonian, like other subspace iteration algorithms.

However, truncation makes things much more complicated since the subspace deviates from the correct occupied subspace after truncation. As a result, the iteration process might not converge, as illustrated in Figure 4. This is a common feature for linear scaling algorithms.
(a) $a = 1000; \sigma = 0.15$.

(b) $a = 100; \sigma = 0.3$.

(c) $a = 10; \sigma = 0.3$.

(d) $a = 10; \sigma = 0.45$.

FIG. 2: Band structure for different values of $\sigma$ and $a$. Choosing $a$ and $\sigma$ appropriately we can control the size of the energy band gap. In figures (a), (b), and (c), we show insulator systems with different size gaps. The gap in the system shown in figure (d) is almost zero, closer to what occurs in a metal.

FIG. 3: Band gap as function of $a$ and $\sigma$. The gap scales like $\sqrt{a/\sigma}$. (a) Fixed $\sigma$, $a$ varies from 0 to 10000. (b) Fixed $a$, $\sigma$ varies from 0.05 to 0.4.

involving truncation. For instance, for the orbital minimization algorithm [27], the truncation step in general increases the energy while the minimization step in general decreases the energy. Hence, even if a variational algorithm is used, it is not guaranteed that it will converge to the minimizer (not to mention other issues, such as the issue of local minimizers in the orbital minimization algorithm [27]).

The general behavior of the trajectory of the energy in the LSI iteration is shown in figure 4. In figure 4(a)
and figure 4(b), the difference between the energy at each step and the accurate result obtained by direct diagonalization is shown in a semi-log plot: In figure 4(b), we also plot the energy before the truncation step. It is obvious that the truncation will increase the energy since the subspace deviates from the occupied subspace after truncation. It can be seen that the error first decays exponentially and then starts to fluctuate around a value that is small but different from the round-off error. In figure 4(c) and figure 4(d), the fluctuations are shown for two different cases. We will call this behavior remanent fluctuation, and we will refer to the error between the numerical solution and the true minimizer (without truncation) remanent error. The remanent error is caused by the truncation of the wave functions.

We will quantify the size of the remanent fluctuation by its variance between iterations $N_0$ and $N_0 + N_s$, defined as

$$ (\Delta E)^2 = \frac{1}{N_s - 1} \sum_{i=N_0+1}^{N_0+N_s} (E_i - \bar{E})^2, \quad (30) $$

where $\bar{E}$ is the mean value of $\{E_i\}$, $i = N_0 + 1, \cdots, N_0 + N_s$. The remanent error in the energy is then the difference between $\bar{E}$ and the true energy of the system. Obviously, $\Delta E$ depends on $N_0$ and $N_s$. In the following, we choose $N_s = 1000$ and $N_0$ sufficiently large such that the iteration has already started to fluctuate before step $N_0$. We calculate the variance $\Delta E$ for different parameters (band gap; cut-off radius $R_c$; order of Chebyshev filter $N_{\text{cheb}}$) to try to understand the effects of these parameters on the dynamical behavior of LSI iteration. The results are collected in Table I. It is observed that the variance decreases when either the cut-off radius, the band gap, or the order of the Chebyshev filter is increased. The effect of the cut-off radius is the most significant in the three cases considered: Indeed, the variance decreases exponentially as the cut-off radius is increased.

We may understand the behavior of the LSI and other related algorithms that include a truncation step by considering a local error estimate. Let us denote by $V_0$ the true occupied subspace of the Hamiltonian and by $V_0$
Now assume that we start from a subspace $V$, results stay within a neighborhood of the true occupied subspace is an attractor for the subspace iteration. With this result, note that, without truncation, the true occupied subspace have inside a small neighborhood of $V$ is determined by the localization property of the Hamiltonian, and hence, $\delta$ is decreased as the band gap is enlarged. Since $d(TL(\cdot, \cdot))$ is a continuous function of the subspace, we have

$$d(TL(V), V) \leq C\delta$$

if subspace $V$ is sufficiently close to $V_0$, say if $d(V, V_0) < M$.

For the filtering step, the error is reduced as in traditional subspace iteration methods:

$$d(F(V_n), V_0) \leq \lambda d(V_n, V_0) = \lambda e_n.$$  

where $\lambda < 1$ is the amplification factor determined by the band gap and order of the Chebyshev polynomial used. Now assume that we start from a subspace $V_n$ that is inside a small neighborhood of $V_0$. Then, by (32), we have

$$e_{n+1} = d(TLF(V_n), V_0) \leq d(TLF(V_n), F(V_n))$$

$$+ d(F(V_n), V_0) \leq \lambda e_n + C\delta.$$  

Here, the estimate (34) is valid when the current subspace is close enough to the true occupied subspace. It is a local result. Note that, without truncation, the true occupied subspace is an attractor for the subspace iteration. With truncation, it is in general not even a fixed point. The estimate (34), however, guarantees that the numerical results stay within a neighborhood of the true occupied space if they come close enough to it. The size of the neighborhood, which is clearly related to the variance of the remanent fluctuation, can be estimated from (34), and it is given by

$$e_\infty \leq C\delta/(1 - \lambda).$$

Note that $\delta$ decreases exponentially as the cut-off radius increases, and $\lambda$ becomes smaller as the band gap or the order of the filter increases. At least qualitatively, this estimate agrees well with the numerical results. We will leave the detailed discussion of the dynamical behavior of LSI and other related algorithms to future publications.

It is clear that, in practice, the LSI iteration should be terminated when the remanent fluctuation is reached. When the specified error tolerance is smaller than the remanent error, the iteration is terminated before the remanent fluctuation is reached. Later, when we talk about the accuracy and convergence of LSI, it should be understood in this sense.

### C. The performance of the LSI for the one-dimensional model

We will focus on three sets of parameter values:

1. $a = 1000$ and $\sigma = 0.15$: The system is an insulator with a fairly large band gap ($E_{\text{gap}} = 323.974$ and $E_{\text{gap}}/E_{\text{width}} = 3193.766$). The potential is quite narrow peaked in this case.

2. $a = 10$ and $\sigma = 0.45$: The system is close to being a metal in this case, with only a tiny gap between the first and second bands ($E_{\text{gap}} = 0.731622$ and $E_{\text{gap}}/E_{\text{width}} = 0.16686$). The external potential is almost flat in this case.

3. $a = 10$ and $\sigma = 0.3$: This is an intermediate situation. There is a small band gap but it is not so large ($E_{\text{gap}} = 3.43881$ and $E_{\text{gap}}/E_{\text{width}} = 0.998883$). The system can be regarded as a semiconductor.

Since the system is periodic, for the exact solution we will use $k$-point sampling with with $64$ $k$-points in the first Brillouin zone $[-\pi, \pi]$. We use a second order finite difference discretization with grid size $1/64$.

First let us study the accuracy of the LSI. We choose a cut-off radius $R_c = 4.0$ (recall that the lattice constant is 1). The results are shown in Table I. In the table, for each case, the energy calculated from direct diagonalization and from the LSI algorithm, the relative error in the energy and the error in the electronic density in the $L^1$ norm are listed. For the insulator (case 1) and the semiconductor (case 3), the results from the LSI algorithm are very accurate. For the metal-like system (case 2), since the gap is quite small, the performance of linear scaling algorithms is expected to deteriorate, but the LSI algorithm still gives results with fairly good accuracy.

<table>
<thead>
<tr>
<th>$a$</th>
<th>$\sigma$</th>
<th>$R_c$</th>
<th>$N_{\text{choo}}$</th>
<th>$\Delta E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.2</td>
<td>1.0</td>
<td>20</td>
<td>$1.2246 \times 10^{-5}$</td>
</tr>
<tr>
<td>10</td>
<td>0.2</td>
<td>2.0</td>
<td>20</td>
<td>$2.1078 \times 10^{-7}$</td>
</tr>
<tr>
<td>10</td>
<td>0.2</td>
<td>3.0</td>
<td>20</td>
<td>$5.8968 \times 10^{-11}$</td>
</tr>
<tr>
<td>10</td>
<td>0.3</td>
<td>1.0</td>
<td>20</td>
<td>$2.3690 \times 10^{-7}$</td>
</tr>
<tr>
<td>10</td>
<td>0.3</td>
<td>2.0</td>
<td>20</td>
<td>$4.6988 \times 10^{-7}$</td>
</tr>
<tr>
<td>10</td>
<td>0.3</td>
<td>3.0</td>
<td>20</td>
<td>$2.0813 \times 10^{-11}$</td>
</tr>
<tr>
<td>10</td>
<td>0.4</td>
<td>1.0</td>
<td>20</td>
<td>$3.3690 \times 10^{-7}$</td>
</tr>
<tr>
<td>10</td>
<td>0.4</td>
<td>2.0</td>
<td>20</td>
<td>$1.8421 \times 10^{-6}$</td>
</tr>
<tr>
<td>10</td>
<td>0.4</td>
<td>3.0</td>
<td>20</td>
<td>$2.6947 \times 10^{-10}$</td>
</tr>
<tr>
<td>25</td>
<td>0.3</td>
<td>1.0</td>
<td>20</td>
<td>$1.2325 \times 10^{-10}$</td>
</tr>
<tr>
<td>25</td>
<td>0.3</td>
<td>1.0</td>
<td>12</td>
<td>$1.2639 \times 10^{-5}$</td>
</tr>
<tr>
<td>25</td>
<td>0.3</td>
<td>1.0</td>
<td>6</td>
<td>$6.3292 \times 10^{-5}$</td>
</tr>
<tr>
<td>25</td>
<td>0.3</td>
<td>1.0</td>
<td>4</td>
<td>$3.3415 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

**TABLE I:** The relationship between the size of remanent fluctuation (quantified by its variance) and the parameters in the LSI iteration.
As noted above, the cut-off radius in the results presented in Table II was fixed at 4.0. We now study the effect that the cut-off radius has on the accuracy of the LSI. Naturally we expect that the larger the cut-off radius used, the more accurate the result will be. As an illustration, we take the parameter values \( a = 100 \) and \( \sigma = 0.3 \) with a cut-off radius that varies from 0.1 to 4.0 lattice constants. The accuracy of the energy is shown in a logarithmic scale in figure 5(a). It is observed that the cut-off radius increases, the error decreases exponentially before reaching the scale of \( 10^{-8} \), which is the stopping criterion used in the iteration for remanent fluctuation. This agrees well with the expected exponential decaying property of the Wannier function for insulators. It is also of interest to see how the computed wave functions change as the cut-off radius is changed. This can be seen in figure 5(b), where we show the wave functions resulting from the LSI with cut-off radii \( R_c = 0.2, 0.6, \) and 1.0. With a small cut-off radius \( R_c = 0.2 \), the result is quite far from the correct result. For \( R_c = 0.6 \) and 1.0, however, the wave functions are only notably different from each other at the edge of the cut-off radius. The wave function obtained using an even larger cut-off radius is almost identical to the one obtained with \( R_c = 1.0 \), and is therefore not included in the figure.

More interesting is the case when the system is metal-like, since this will help us understand how the LSI fails. Therefore we will focus next on the metal-like case when \( a = 10 \) and \( \sigma = 0.45 \). In figure 6(a) the error in energy is shown in a log-log plot as the cut-off radius is varied from 2.0 to 13.0. It is observed that the error decays algebraically as the cut-off radius increases [37]. The resulting wave functions corresponding to cut-off radii 3, 6 and 9 are shown in figure 6. The wave function expands and becomes closer to the correct solution as the cut-off radius grows. Note also that a larger cut-off radius is required to obtain accurate results for metal-like systems than for insulator systems (\( R_c = 1.0 \) is enough for the previous insulator case).

### IV. NUMERICAL EXAMPLES: VALIDATION AND LINEAR SCALING

We have implemented the LSI algorithm for Kohn-Sham DFT in three dimensions. To achieve linear scaling, it is important to use a sparse representation for the wave functions, as well as the pseudopotential. Both the wave functions and the pseudopotential components are compactly supported, and therefore once the grid size is fixed, each array contains only a small fraction of grid values.

Consider the domain \( \Omega = [0, L] \times [0, D] \times [0, H] \), containing the support of all the wave functions. This domain contains also the support of all the pseudopotentials components (local and nonlocal), as well as the support of the atomic function. We discretize \( \Omega \) using a uniform mesh with grid sizes \( \Delta x = L/n_x, \Delta y = D/n_y, \) and \( \Delta z = H/n_z \). We define the values of the wave functions and the density at the center of the cells: \( \psi_{i,j,k} \approx \psi(x_i, y_j, z_k) \), where

\[
\begin{align*}
  x_i &= \left(i - \frac{1}{2}\right) \Delta x, \quad i = 0, \ldots, n_x + 1, \\
  y_j &= \left(j - \frac{1}{2}\right) \Delta y, \quad j = 0, \ldots, n_y + 1, \\
  z_k &= \left(k - \frac{1}{2}\right) \Delta z, \quad k = 0, \ldots, n_z + 1.
\end{align*}
\]

The points with indices \( 0 \) or \( n_x + 1, n_y + 1, \) and \( n_z + 1 \) are ghost cells outside the computational domain, and are used only to impose the boundary conditions. Since the wave functions are zero on the boundary, we define the ghost values by reflection, which on the \( x = 0 \) boundary would be: \( \psi_{0,j,k} = -\psi_{i,j,k} \). A similar expression is used on the other boundaries.

The Laplacian is discretized using second order centered differences with the standard seven point stencil. If we define

\[
\begin{align*}
  \delta_{xx} u_{i,j,k} &= u_{i+1,j,k} - 2u_{i,j,k} + u_{i-1,j,k}, \\
  \delta_{yy} u_{i,j,k} &= u_{i,j+1,k} - 2u_{i,j,k} + u_{i,j-1,k}, \\
  \delta_{zz} u_{i,j,k} &= u_{i,j,k+1} - 2u_{i,j,k} + u_{i,j,k-1},
\end{align*}
\]

the Laplacian is

\[
\Delta u_{i,j,k} = \delta_{xx} u_{i,j,k} + \delta_{yy} u_{i,j,k} + \delta_{zz} u_{i,j,k}.
\]

The Coulomb term is approximated as a discrete convolution, and this convolution is computed using the Fast Fourier Transform (FFT) with zero padding, as described in [11] [29].

To validate our code, we have compared the result obtained with our algorithm with the results obtained using the package PARSEC [29] [31]. We considered two alkane chains \( \text{CH}_3(\text{CH}_2)_n \text{CH}_3 \) with \( n = 0 \) (Ethane) and with \( n = 10 \). The geometric structure of the alkane was determined using the geometry optimization option in the Gaussian package [32], using the Gaussian-type orbitals STO-3G. The molecule so obtained for \( n = 10 \) is plotted in figure 7(a) to illustrate the geometric structure. Although the molecule is a linear chain, there is a nontrivial three-dimensional structure due to the Hydrogen atoms.

We used the Troullier-Martins pseudopotentials in the Kleinman-Bylander form [16] [17]. For Hydrogen, the
pseudopotential has one local component for the 1s orbital, and no nonlocal components; for Carbon, however, we need to consider the 2s and 2p nonlocal components of the pseudopotential. We choose the 2p pseudopotential to be the local component, in order to avoid using spherical harmonics. The nonlocal pseudopotential is therefore the 2s component (so \( l_{\max} = 0 \) in (4)). We generate the pseudopotentials using the code developed by Paolo Gianozzi and his collaborators, available online at [33].

For the exchange and correlation term in the energy, we consider the homogeneous electron gas approximation of Ceperley and Alder [34], as parameterized by Perdew and Zunger [35].

The pseudopotential operator acts on the wave function \( \psi \) in the following way:

\[
\hat{V}_{PS} \psi(x) = \sum_{j=1}^{N} \left( V_{local}(x - R_j) \psi(x) + \alpha_j \beta^{2p}_{00}(x - R_j) \right),
\]

(38)

where we have defined

\[
\alpha_j = \int_{\mathbb{R}^3} \beta^{2p}_{00}(y - R_j) \psi(y) \, dy, \quad j = 1, \ldots, N_a.
\]

(39)

Both \( \beta^{2p}_{00}(\cdot - R_j) \) and \( \psi \) are compactly supported, so the integrals only need to be carried out if the supports of \( \beta^{2p}_{00}(\cdot - R_j) \) and \( \psi \) intersect. The integral is approximated using the midpoint rule [36]:

\[
\alpha_j = \Delta V \sum_{kln} \beta^{2p}_{00}(y_{kln} - R_j) \psi_{kln},
\]

(40)

where \( \Delta V = \Delta x \Delta y \Delta z \) is the volume of each computational cell. The wave functions are centered at the center of the chemical bonds, and initially they are chosen to be Gaussians. We also considered the case of random initial wave functions, and the code converged as well, although the number of iterations was higher. The self-consistent iteration was terminated when the difference between two consecutive electron densities, measured in the \( L_2 \)-norm,
was less than $10^{-5}$. In principle, the filtering and localization steps must be performed until the linear iteration has converged, as described in algorithm\textsuperscript{[1]}. In practice, however, we only perform these steps once before updating the electron density and the effective potential.

For validation purposes, we have compared the energies obtained with the LSI code and the corresponding energies obtained with PARSEC. The result of the comparison is shown in table\textsuperscript{[III]}\ As can be seen, the difference is of the order of $2 - 3\%$. This discrepancy is due to the differences in the formulation and implementation of the two codes, such as domain size, discretization, evaluation of the Coulomb potential, etc., and are of no consequence for the purposes of this article. The PARSEC code was used only as a reference to validate our own. The electron density obtained for $N = 38$ atoms is shown in figure\textsuperscript{[V]}.

<table>
<thead>
<tr>
<th>Comparison with Parsec</th>
<th>( C_2H_6 ); ( N = 8 )</th>
<th>( CH_3(CH_2)_9CH_3; \ N = 38 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E_{\text{LSI}} )</td>
<td>( E_{\text{con}} )</td>
<td>( E_{\text{KS}} )</td>
</tr>
<tr>
<td>( -5.62055847 )</td>
<td>( -22.15732546 )</td>
<td>( -14.38217632 )</td>
</tr>
<tr>
<td>( E_{\text{PARSEC}} )</td>
<td>( -5.35870240 )</td>
<td>( -22.35123917 )</td>
</tr>
<tr>
<td>( E_{\text{diff}} )</td>
<td>( 1.58% )</td>
<td>( 0.87% )</td>
</tr>
</tbody>
</table>

TABLE III: Comparison between our code and PARSEC. The ionic energy \( E_{\text{con}} \) includes the Coulomb and pseudopotential energies. The results of the LSI are in good agreement with the results from PARSEC.

To illustrate the linear scaling behavior of our code, we have collected the timings for the filtering, localization, inversion of the Gram matrix, and computation of the Fermi energy for alkane chains with sizes ranging between 38 and 542 atoms. The results are shown in figure\textsuperscript{[VI]}.

We have also computed the error in the approximation of \( S^{-1} \) in the Frobenius norm. The results are shown in table\textsuperscript{[IV]}\ The tolerance was set to \( \varepsilon = 10^{-5} \). The error appears to be independent of the size of the molecule. A reduction in the error can be achieved by simply reducing the tolerance.

\[ \| S^{-1} - S^{-1} \|_F \]

\[ \begin{array}{c|c|c|c|c|c|c}
 N & \| S^{-1} - S^{-1} \|_F \\
\hline
 38 & 2.850907 \times 10^{-5} \\
 181 & 2.885210 \times 10^{-5} \\
 217 & 2.854986 \times 10^{-5} \\
 253 & 2.851860 \times 10^{-5} \\
 289 & 2.850546 \times 10^{-5} \\
 325 & 2.839061 \times 10^{-5} \\
 361 & 3.122993 \times 10^{-5} \\
 397 & 3.152983 \times 10^{-5} \\
 433 & 3.054400 \times 10^{-5} \\
 469 & 3.226956 \times 10^{-5} \\
 505 & 3.301841 \times 10^{-5} \\
 541 & 3.279986 \times 10^{-5} \\
 577 & 3.348167 \times 10^{-5} \\
 649 & 3.298259 \times 10^{-5} \\
\end{array} \]

TABLE IV: Error in the computation of the inverse of the Gram matrix as a function of molecular size. The errors appear to be independent of the size of the system in this case.

We have also seen that the added difficulty associated with the non-orthogonal Gram matrix of the non-orthogonal wave functions can be overcome rather easily.

We have presented and analyzed a controllable one-dimensional model with which we can reproduce the main characteristics of insulating and metallic systems. This model has been used to illustrate the convergence properties of the LSI, and we plan to use it as a benchmark for the analysis of other methodologies.

We have also implemented the LSI in a realistic three-dimensional setting and tested its accuracy, as well as its linear scaling properties. All together, these demonstrate that the LSI is a very promising tool for the electronic structure analysis of insulators.

VI. ACKNOWLEDGMENTS

The authors would like to thank Weiguo Gao and Emil Prodan for interesting discussions. The work of Garcia-Cervera is supported by and NSF CAREER award (DMSS-0645766). The work of E, Lu, and Xuan is supported in part by ONR grant N00014-01-1-0674, DOE grant DOE DE-FG02-03ER25587, NSF grants DMS-0407866 and DMR-0611562.
FIG. 7: Electron density of an alkane chain obtained with the LSI algorithm. (a) Atomic configuration, obtained using the package GAUSSIAN \cite{ref32}. (b) Electron density.

FIG. 8: Timings obtained with the LSI code. Linear scaling is observed.

\[ \text{FIG. 7: Electron density of an alkane chain obtained with the LSI algorithm. (a) Atomic configuration, obtained using the package GAUSSIAN [32]. (b) Electron density.} \]

\[ \text{FIG. 8: Timings obtained with the LSI code. Linear scaling is observed.} \]


http://www.fisica.uniud.it/~giannozz/.


[37] Strictly speaking, the system still has a tiny gap, but the exponent of exponential decay is so small that the algebraic behavior is dominant.